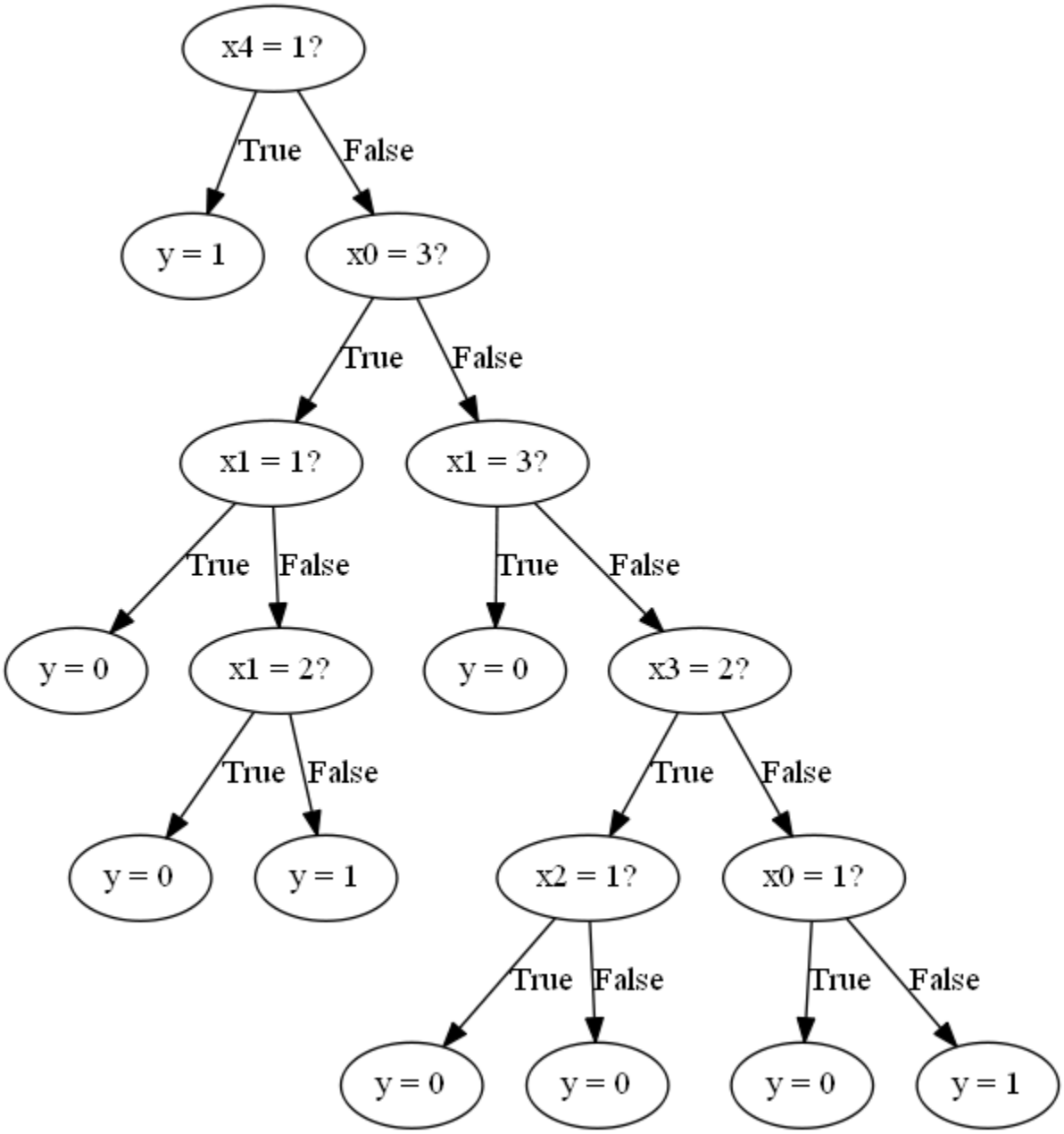
Name : Lakshmeesha Suresh Shetty

Net-ID : LSS180005

Machine Learning Assignment 2

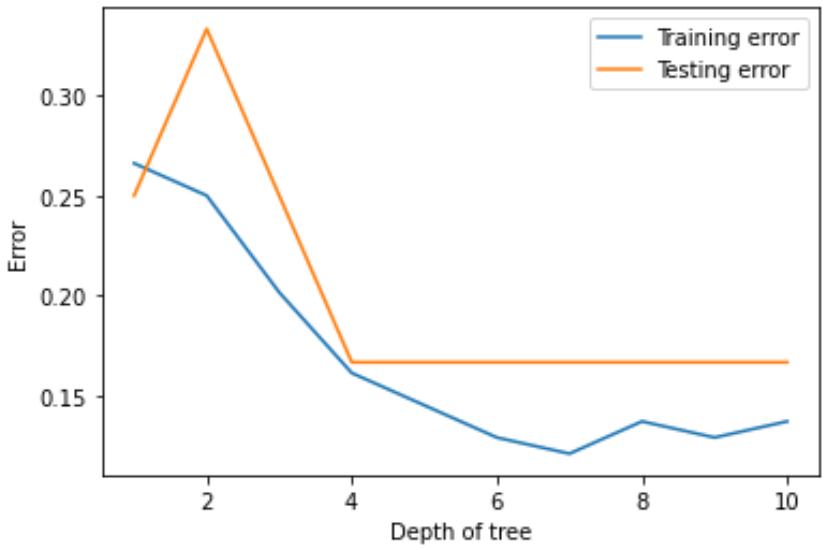
a => Decision Tree for depth = 5



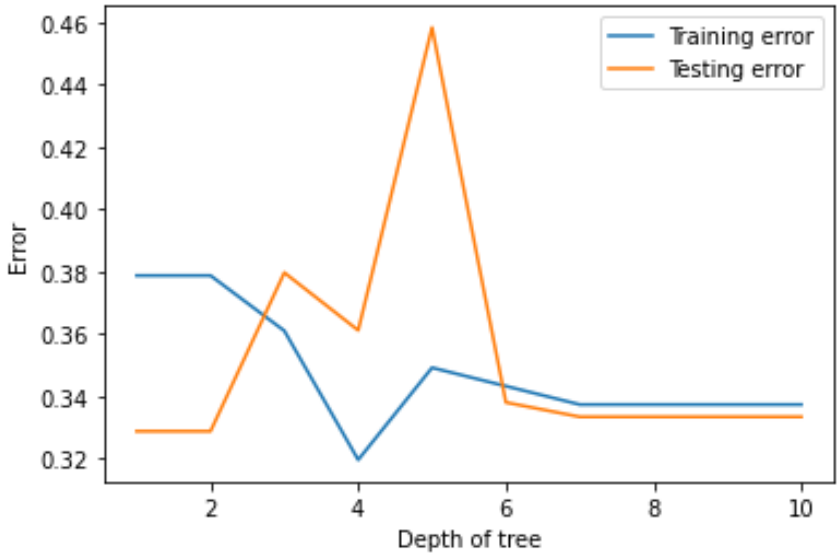
b =>

The following are the plots obtained for each of the monk's dataset for training and testing errors:

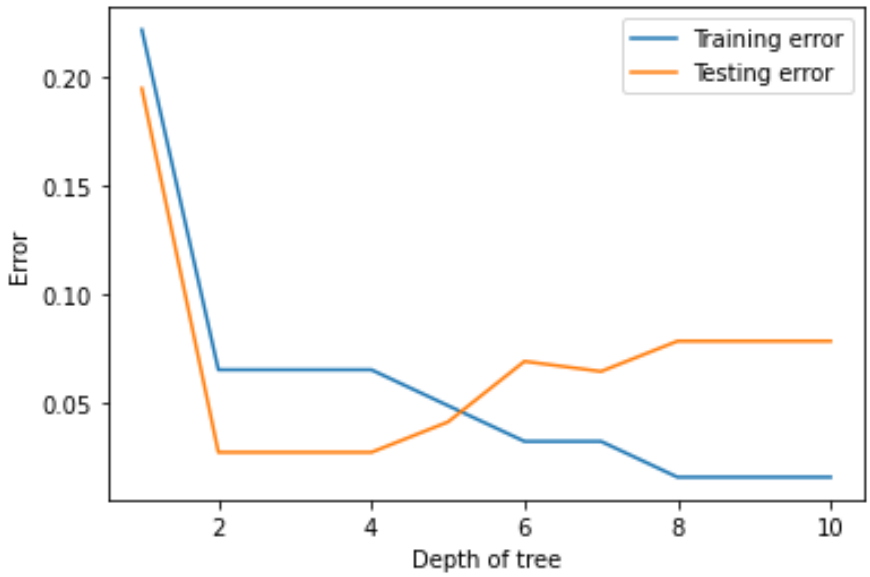
Monks-1 Dataset



Monks-2 Dataset



Monks-3 Dataset

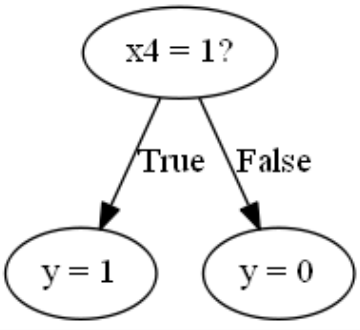


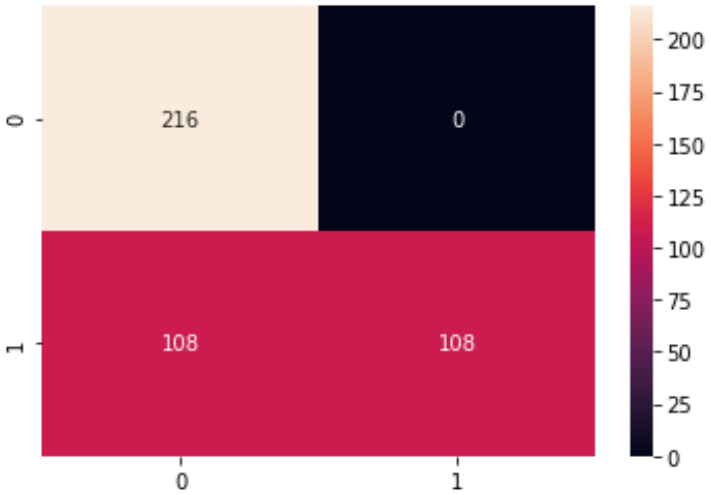
In each of the following plots we can see that for initial lower depths of tree, our model under-fits in performance but as we increase the depth of tree, we see a decrease in both training and testing error until they start overfitting. So, with a depth of tree between 6 to 8 would be better for our model.

c =>

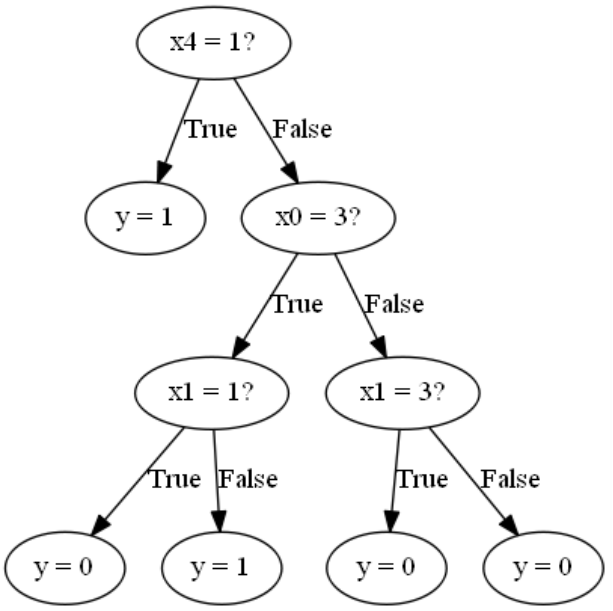
The Decision tree and confusion matrix for our algorithm for depth = 1,3 and 5 are as follows:

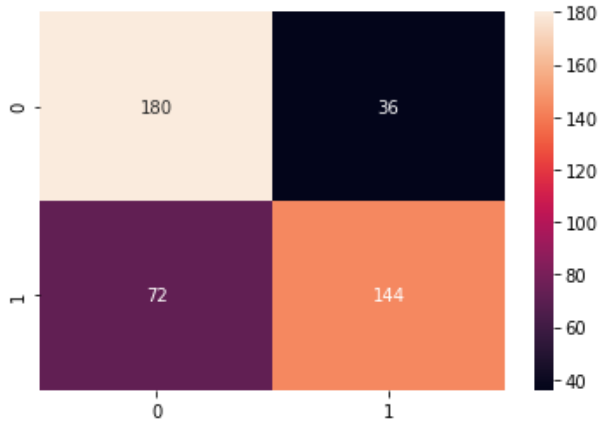
1. Decision Tree and confusion matrix of depth = 1



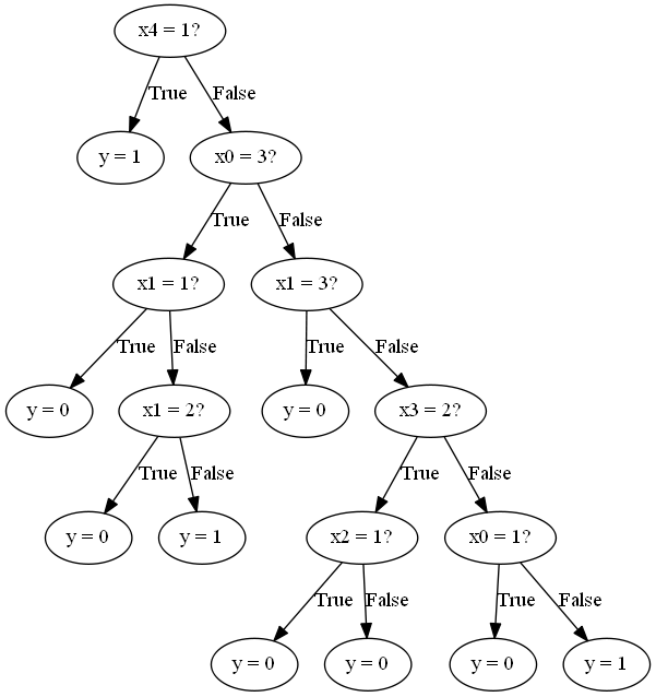


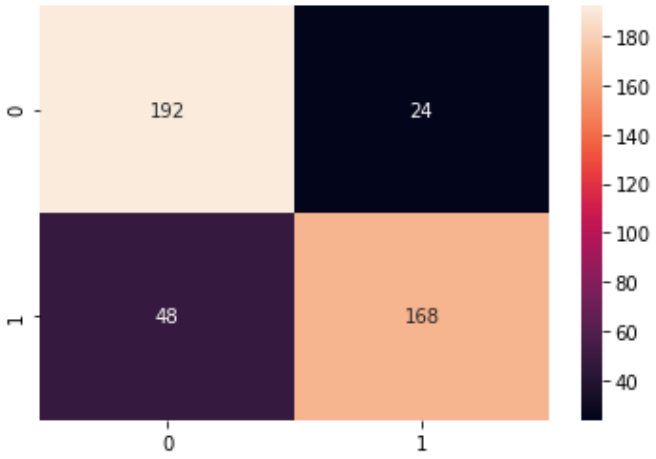
2. Decision Tree and confusion matrix of depth = 3





3. Decision Tree and confusion matrix of depth = 5

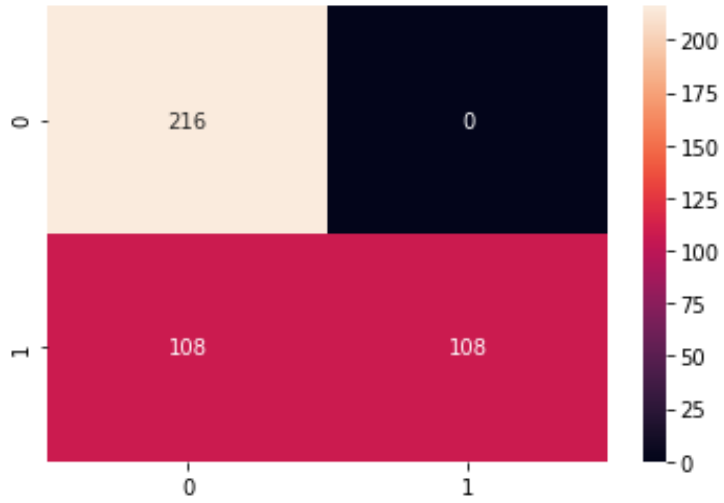




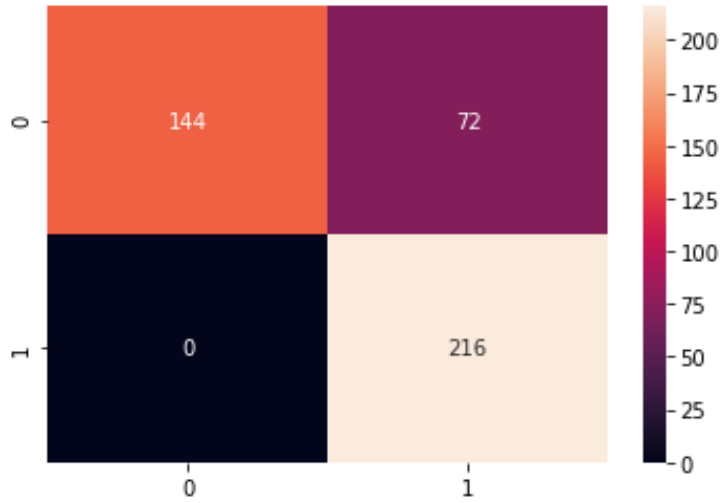
d =>

The Confusion matrix for scikit-learn's DecisionTreeClassifier algorithm for depth = 1,3 and 5 are as follows:

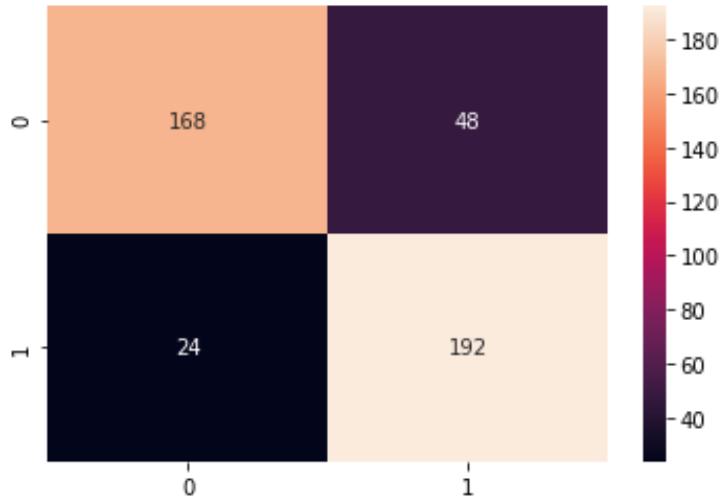
1. Confusion matrix of depth = 1



2. Confusion matrix of depth = 3



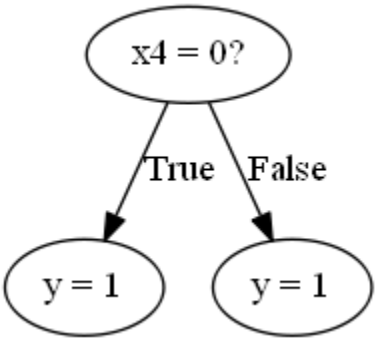
3. Confusion matrix of depth = 5

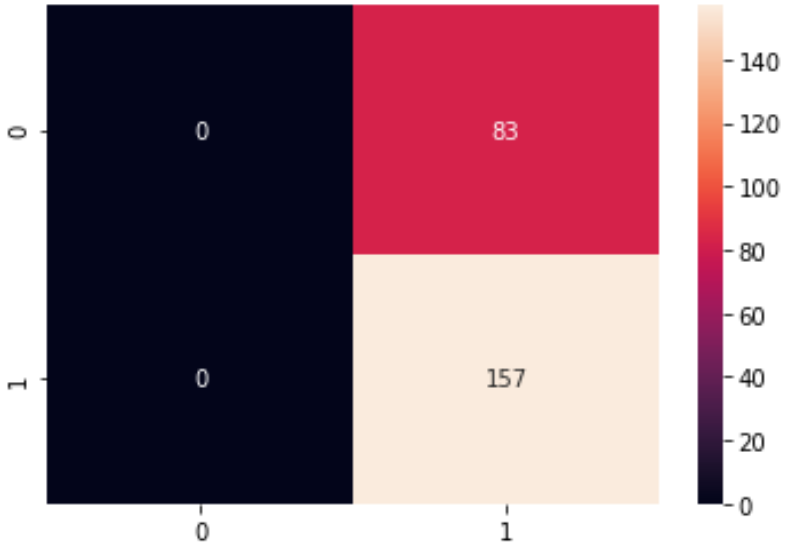


e =>

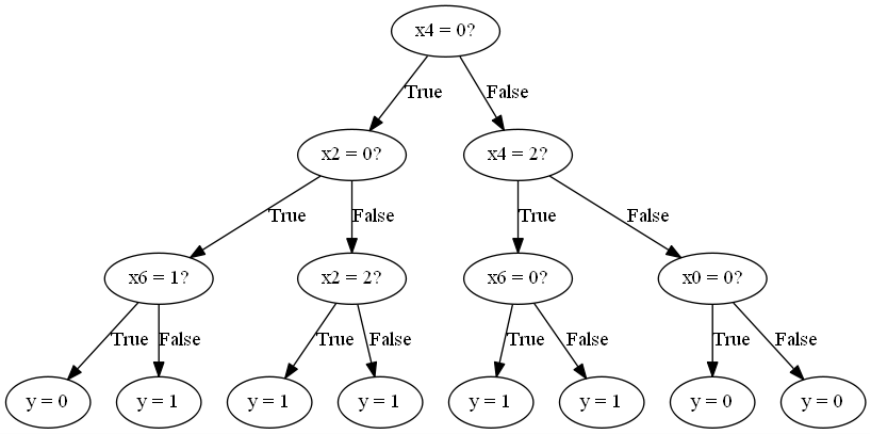
The Decision tree and confusion matrix for our algorithm for a different dataset is as follows:

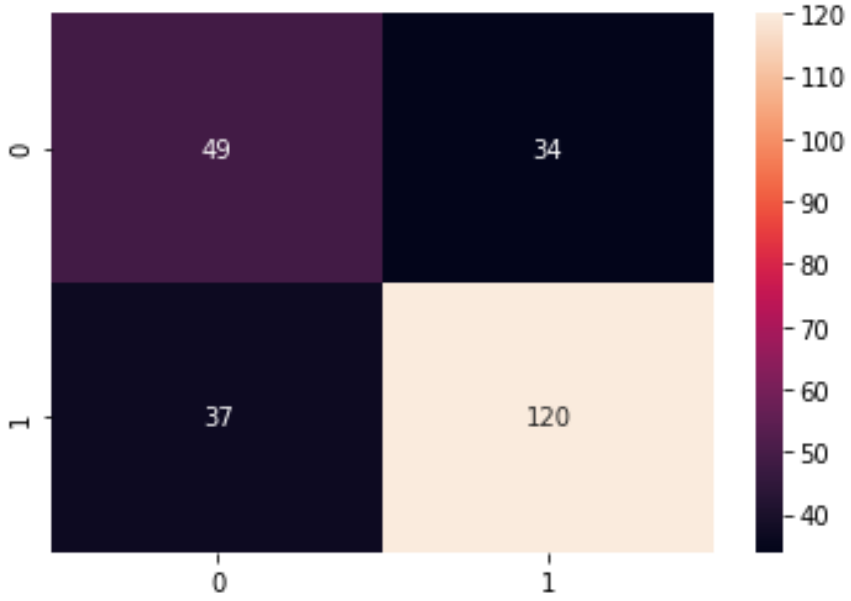
1. Decision tree and confusion matrix for depth = 1



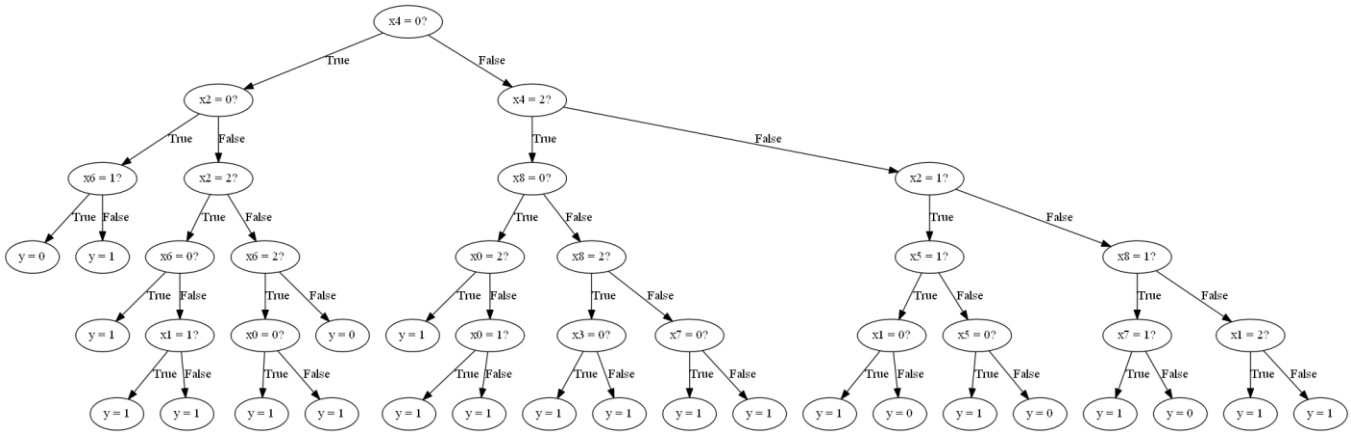


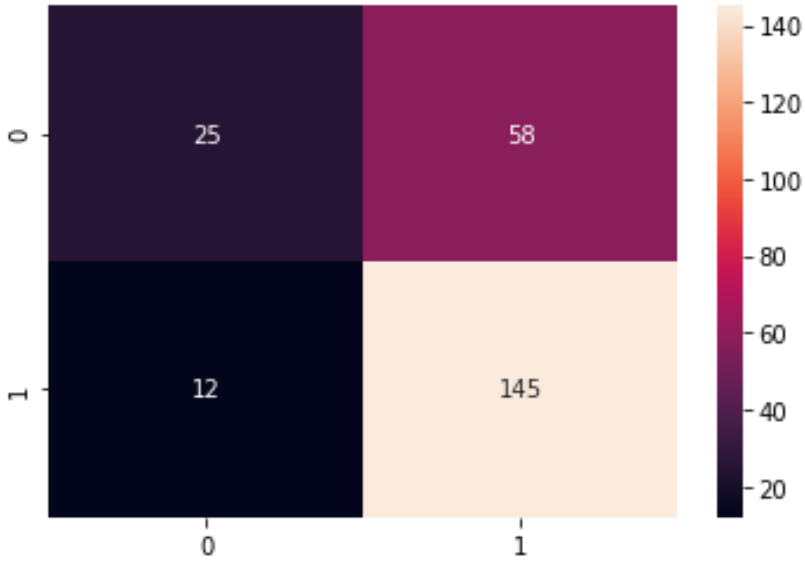
2. Decision tree and confusion matrix for depth = 3





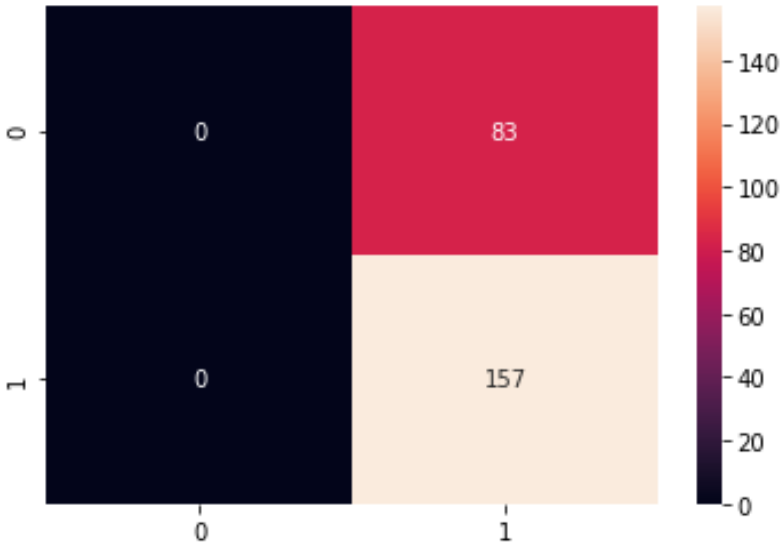
3. Decision tree and confusion matrix of depth = 5



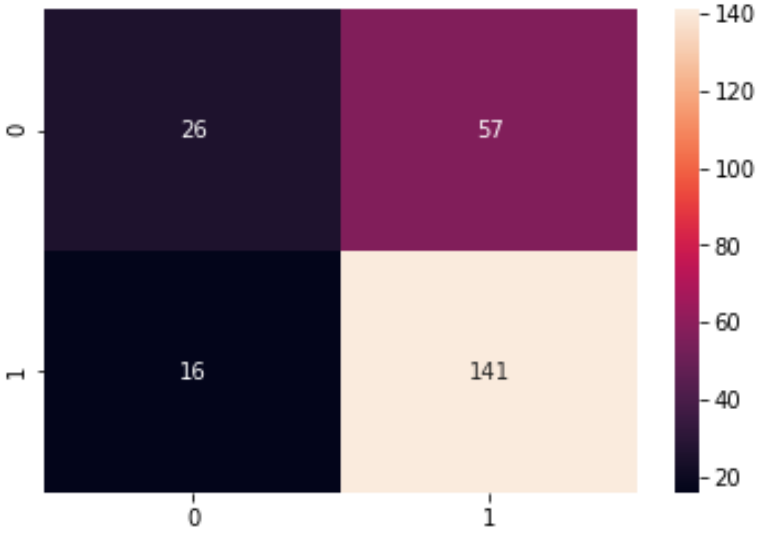


The Decision tree and confusion matrix for scikit-learn's DecisionTreeClassifier are as follows:

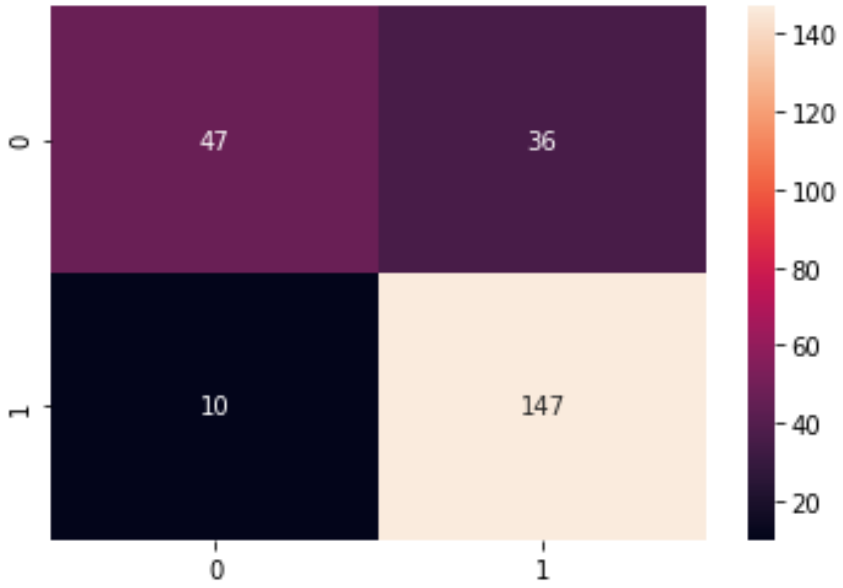
1. Confusion matrix for depth = 1



2. Confusion matrix for depth = 3



3. Confusion matrix for depth = 5



Summary:

The following report shows that our ID3 algorithm implementation is only far from scikit-learn's DecisionTreeClassifier by a small margin in terms of accuracy of prediction. According to the results for Depth = 5 decision tree for both the algorithm, ID3 had a test error of 29.17% while DecisionTreeClassifier had a test error of 21.25%. So our implementation of Decision tree is not so far from scikit-learn's algorithm in terms of accuracy.